

**TABLE OF CONTENTS**

<b>CHAPTER</b>	<b>TITLE</b>	<b>PAGES</b>
	<b>DECLARATION</b>	<b>ii</b>
	<b>DEDICATION</b>	<b>iii</b>
	<b>ACKNOWLEDGEMENT</b>	<b>iv</b>
	<b>ABSTRACT</b>	<b>v</b>
	<b>ABSTRAK</b>	<b>vi</b>
	<b>TABLE OF CONTENTS</b>	<b>vii</b>
	<b>LIST OF TABLES</b>	<b>xiv</b>
	<b>LIST OF FIGURES</b>	<b>xix</b>
	<b>LIST OF ABBREVIATION</b>	<b>xxi</b>
	<b>LIST OF APPENDICES</b>	<b>xx</b>
<b>1</b>	<b>INTRODUCTION</b>	<b>1</b>
	1.1 Background of the Problem	2
	1.2 Motivation of Research	4
	1.3 Problem Statement	7
	1.4 Goal	8
	1.5 Objectives	8
	1.6 Scope	9
	1.7 Significance of the research	9
	1.8 Research Contributions	11

1.9	Thesis Organization	12
<b>2</b>	<b>LITERATURE REVIEW</b>	<b>14</b>
2.1	Speaker Identification	16
2.2	Speaker Verification	18
2.3	Front-end Processing/Feature Extraction	19
2.3.1	Mel Frequency Cepstral Coefficients (MFCC)	22
2.3.1.1	Frame Blocking	23
2.3.1.2	Windowing	24
2.3.1.3	Fast Fourier Transform	24
2.3.1.4	Mel-Frequency Wrapping	25
2.3.1.5	Cepstrum	25
2.4	Back-end Processing/ Pattern Classification	26
2.4.1	Evolution of Pattern Classification Technique in Speaker Identification	28
2.4.2	Dynamic Time Warping (DTW)	30
2.4.2.1	DTW in Speaker Identification	31
2.4.3	Hidden Markov models (HMM)	32
2.4.3.1	HMM in Speaker Identification	33
2.4.4	Vector Quantization (VQ)	36
2.4.4.1	VQ in Speaker Identification	37
2.4.5	Gaussian Mixture Models Classification Model	39
2.4.5.1	GMM in Speaker Identification	40
2.4.6	Neural Networks for Classification	41
2.4.6.1	Neural Networks in Speaker Identification	43
2.4.7	Support Vector Machines (SVM)	44
2.4.7.1	SVM in Speaker Identification	45
2.5	Evaluation on Several Pattern Classification Techniques	46

2.6	Comparison on Several Pattern Classification Approaches	51
2.7	Recent Work Progress on GMM in Speaker Identification	53
2.8	Recent Work on Hybrid Modeling	56
2.9	Summary	59
<b>3</b>	<b>RESEARCH METHODOLOGY</b>	<b>60</b>
3.1	The Incentive of Hybrid VQ/GMM Model	60
3.2	Recent Works on Hybrid VQ/GMM Modeling	63
3.2.1	Constrain of Hybrid VQ/GMM Modeling	66
3.2.2	Review of VQ Problem Solutions	66
3.2.3	Distributed VQ Training	68
3.3	Operational Framework of Research Model	69
3.4	Design Methodology	71
3.4.1	Data Collection (Speech Corpus)	73
3.4.2	Distributed Data Pre-processing Phase	75
3.4.2.1	Rules of Decision Tree – Pitch Analysis	76
3.4.2.2	Subgroup Data Distribution	76
3.4.2.3	Level of Decision Tree	79
3.4.2.4	Process Flow of Data Distribution	80
3.4.3	Distributed VQ Training Model	81
3.4.3.1	LBG Algorithm	82
3.4.4	Applying Distributed VQ Clustering as Pre-classifier for GMM	83
3.4.4.1	Data Selection for Initial Test Set	85
3.4.5	Speaker Identification using GMM Likelihood Ratio Algorithm	87
3.5	Summary	89

<b>4</b>	<b>HYBRID DISTRIBUTED VQ/GMM MODELING</b>	<b>90</b>
4.1	Overview of the Design Framework	91
4.2	Hybrid Training Phase I: Distributed Data Pre-Processing Phase	91
4.2.1	Pitch Analysis Process	92
4.2.2	Threshold Analysis	95
4.2.3	Distributed Data Process	95
4.3	Hybrid Training Phase II: Distributed VQ Training Model	97
4.4	Hybrid Testing Phase I: Distributed VQ Classification Model	99
4.4.1	Nearest Neighbour Search and Distance Calculation	101
4.4.2	Best Speaker Model Selection	101
4.5	Hybrid Testing Phase II: GMM Identification Model	102
4.6	Summary	104
<b>5</b>	<b>RESULTS AND ANALYSIS</b>	<b>105</b>
5.1	Evaluation Measures	106
5.2	Experimental Setup and Condition	107
5.3	Amount of Data Chosen in Experiments	108
5.4	Experiment I: Conventional Baseline System	108
5.4.1	GMM Baseline Speaker Identification Performance	109
5.4.2	VQ Baseline Speaker Identification Performance	112
5.4.3	Comparison between VQ and GMM Performance	114
5.5	Experiment II: Hybrid VQ/GMM Modeling	116
5.5.1	Performance of Hybrid VQ/GMM Model without Distributed Training	117

5.5.2	Computational Time for Hybrid VQ/GMM Model (without Distributed Training)	120
5.5.3	The Performance of Hybrid Distributed VQ/GMM Model	121
5.5.4	Computational Time for Hybrid Distributed VQ/GMM Model	124
5.6	Experiment III: Evaluation on Hybrid Model Vs Baseline Model	124
5.6.1	Evaluation on Processing Time	125
5.6.2	Evaluation on Accuracy Rates	126
5.7	Experiment IV: Evaluation on Other Hybrid VQ/GMM Model	127
5.7.1	VQ Pre-classifier for Gaussian Selection Model	128
5.7.2	Evaluation on VQ Pre-classifier for Gaussian Selection Model	131
5.7.3	LBG Training for GMM Model	133
5.7.4	Evaluation on LBG Training for GMM Model	135
5.8	Summary	137
<b>6</b>	<b>CONCLUSION AND FUTURE WORK</b>	<b>138</b>
6.1	Summary	138
6.2	Future Works	140
	<b>REFERENCES</b>	<b>143</b>
	<b>APPENDIX A-C</b>	<b>162-167</b>

## LIST OF TABLES

TABLE NO.	TITLE	PAGE
2.1	Comparison on several pattern classification approaches	51
3.1	Dialect distribution of speakers	74
5.1	Time use for training and testing for GMM (seconds)	111
5.2	Time use for training and testing for VQ (seconds)	114
5.3	Time use for training /testing for VQ and GMM on full TIMIT data	115
5.4	Computational time use for full TIMIT data training /testing for VQ, GMM and hybrid VQ/GMM model (without distributed data training)	120
5.5	Computational time use for full TIMIT data training /testing on 2 hybrid VQ/GMM model	124
5.6	Processing time for VQ, GMM and hybrid distribute VQ/GMM on full set data	125
5.7	Time used for training /testing on 4 difference models	133
5.8	Comparison on 4 difference models (processing time)	137

## LIST OF FIGURES

FIGURE NO.	TITLE	PAGE
2.1	The basic structure for speaker identification	17
2.2	The basic structure for speaker verification	18
2.3	Mel Frequency Cepstral Coefficients Process	23
2.4	DTW model	31
2.5	HMM finite state generators	35
2.6	Conceptual diagram illustrating Vector Quantization codebook formation	38
2.7	A multilayer perceptron	42
2.8	Steps for Binary Linear Decision Boundary	45
3.1	Hybrid distributed VQ/GMM speaker identification process	69
3.2	Design methodology for hybrid distributed VQ/GMM model	72
3.3	Rules for decision tree to distribute speaker dataset	79
3.4	Flow diagram of the LBG algorithm	83
3.5	Centroids and its code vectors	84
3.6	A process flow of GMM approach in training and testing phase for speaker identification system	88
4.1	Hybrid distributed VQ/GMM modeling - Training phase I	92
4.2	Pseudo code for low pass filtering	93
4.3	Pseudo code of decision tree in distributing speakers' data	96

4.4	Hybrid distributed VQ/GMM modeling - Training Phase II	97
4.5	Distributed VQ algorithm for training data	98
4.6	Hybrid distributed VQ/GMM modeling - Testing Phase I	99
4.7	Pseudo code for decision tree to assign speakers' data into groups	100
4.8	Selection of 6 best speaker models	102
4.9	Hybrid distributed VQ/GMM modeling - Testing Phase II	103
4.10	Algorithm of GMM test on 6 initial data	103
5.1	GMM baseline speaker identification performance	110
5.2	VQ baseline speaker identification performance	112
5.3	Comparisons of VQ and GMM baseline speaker identification performance	115
5.4	Performance of hybrid VQ/GMM model (without distributed data training)	119
5.5	Performance of hybrid distributed VQ/GMM model	123
5.6	The Performance of 3 type of pattern classification models	127
5.7	Structure of VQ pre-classifier for Gaussian selection model	129
5.8	Performance of VQ pre-classifier for Gaussian selection	131
5.9	Comparison of hybrid distributed VQ/GMM model with VQ pre-classifier for Gaussian selection model	132
5.10	Performance of LBG Training for GMM model	135
5.11	Comparison of hybrid distributed VQ/GMM with LBG training for GMM model	136



## LIST OF ABBREVIATIONS

VQ	-	Vector Quantization
GMM	-	Gaussian Mixture Model
DTW	-	Dynamic Time Warping
HMM	-	Hidden Markov Models
FFT	-	Fast Fourier Transform
LPC	-	Linear Prediction Coding
SVM	-	Support Vector Machine
MFCC	-	Mel-Frequency Cepstrum Coefficients
PLP	-	Perceptual Linear Predictive
MAP	-	Maximum A Posterior
VQPGS	-	VQ Pre-classifier for Gaussian Selection
LBG-GMM	-	LBG Training for GMM Model
ZCR	-	Zero Crossing Rate
EM	-	Expectation Maximization
ML	-	Maximum likelihood

**LIST OF APPENDICES**

<b>APPENDIX</b>	<b>TITLE</b>	<b>PAGE</b>
A	List of Publication Related with the Research	163
B	The Result of VQ Pre-classifier for Gaussian Selection Model	167
C	The Result of LBG Training for GMM Model Compare with GMM Baseline Model	168